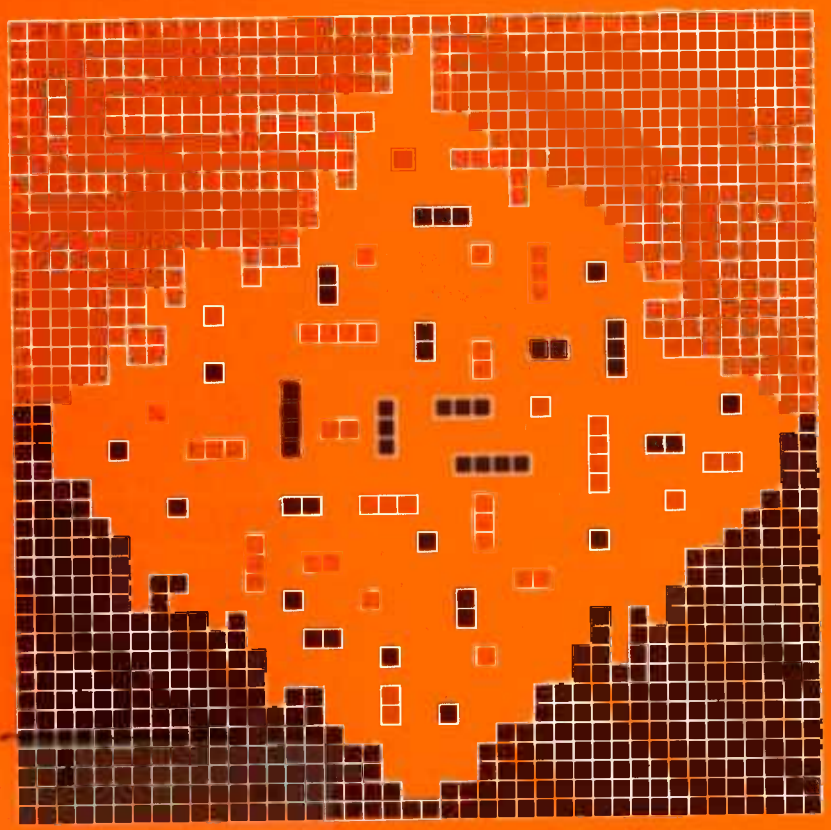
IDRC-156e

International Cooperative Information Systems

Proceedings of a seminar held in Vienna, Austria, 9-13 July 1979



y the International Development Research Centre
eration of the International Atomic Energy Agency
Secretariat for the United Nations Conference
Science and Technology for Development

The International Development Research Centre is a public corporation created by the Parliament of Canada in 1970 to support research designed to adapt science and technology to the needs of developing countries. The Centre's activity is concentrated in five sectors: agriculture, food and nutrition sciences; health sciences; information sciences; social sciences; and communications. IDRC is financed solely by the Parliament of Canada; its policies, however, are set by an international Board of Governors. The Centre's headquarters are in Ottawa, Canada. Regional offices are located in Africa, Asia, Latin America, and the Middle East.

Postal Address: Box 8500, Ottawa, Canada K1G 3H9
Head Office: 60 Queen Street, Ottawa

IDRC, Ottawa CA
IAEA, Vienna AT
UN. Conference on Science and Technology for Development, New York US
IDRC-156e

International Cooperative Information Systems : proceedings of a seminar held in Vienna, Austria, 9–13 July 1979. Ottawa, Ont., IDRC, 1980. 111 p. : ill.

/IDRC publication/, /international cooperation/, /scientific cooperation/, /scientific information/, /technical information/, /information exchange/, /information systems/ — /information services/, /information network/, /data banks/, /AGRIS/, /DEVSIS/, /INIS/, /ISIS/, /UNIDO/, /UNISIST/, /IDRC mentioned/.

UDC: 061.88

ISBN: 0-88936-252-1

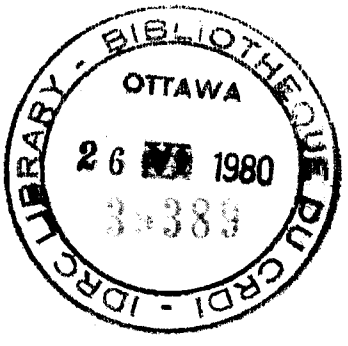
Microfiche edition available

IDRC-156e

International Cooperative Information Systems

Proceedings of a seminar held in Vienna, Austria, 9–13 July 1979

Organized by the
International Development Research Centre
with the cooperation of the
International Atomic Energy Agency
and the
Secretariat for the United Nations Conference
on Science and Technology for Development



Contents

International Cooperative Information Systems

John E. Woolston

Information Sciences Division, International Development Research Centre, Ottawa, Canada

I have been asked to introduce the topic of international cooperative information systems. I do not want to go into any detailed description of, for example, INIS; Dr Pryor will be doing that. Nor of AGRIS; Dr Cazacu will be doing that. What I am going to do today is to talk about the principles that underlie these systems and the way in which they have historically come on the scene, what they offer to the world, and in particular to the developing countries of the world, in the field of information that is needed for policymaking and for decision-making, first, as applied to science and technology, but also as applied more generally to economic and social development.

Let me take the four words in the title of the seminar, "international cooperative information systems," begin with the last one "systems," define it, and work backward adding each of the qualifying terms as I go. First, "system." Perhaps of all the terms, this is the most difficult to handle. The word "system" is one of the most overworked words in the English language. We all use the term for many different things, and the dictionary is of little help in defining "system" in the context in which we are using it.

I like to think of a system as being an organized set of operations. If one is talking about artificial systems as opposed to natural systems, then the organization of the set of operations implies a design, a design that can be modified as time goes on in the light of experience. But the organization did not just happen; it did not come into being spontaneously; it was designed. Systems have an input and an output. And, therefore, since they have an input and an output, their efficiency can be measured. How much work is needed to provide the input? How much useful work is achieved by the output? And the ratio of these, whether expressed quantitatively or in some more subjective way, is a measure of the efficiency of the system.

There are many artificial systems. On Saturday night, I left Canada and entrusted my baggage to the airline baggage delivery system. Alas, my baggage was not at the airport when I got to Vienna. That system disappointed me. All of us expect a system to deliver, and we expect it to behave for us in a rather straightforward way. One of the things we most look for from systems is simplicity from the point of view of the user, accepting that for the designer and for the operator it is indeed something quite complex.

I would now like to qualify the word "system" by the word "information" and talk about "information systems." An "information system" is something that delivers information that is useful to the consumer, to the user of the system.

This text is an edited transcript of a tape-recording of Mr Woolston's remarks.

But, of course, no information system generates information spontaneously. It can only give to the user what has previously been given to it as input. What it delivers to the user is in fact a subset of what it has been given.

An information system works well if, when given a subset from the total mass of information that is within the system, it responds effectively to the selection criteria put to it by the user. A user can come to the system and say, "Give me the information that is in the system and that deals with a particular subject": that is applying *subject* as a parameter of selection. One could also go to the system and say "Give me all the information that deals with a particular country or a particular region of the world": that is applying a *geographic* parameter to the selection of information from the system. One can ask the system to give information that originally came from particular *sources*, from particular *institutions*, or particular *individuals*. Or one can ask an information system for information of a particular *type*: "Give me the information that has statistics within it," for example. So by specifying these different parameters, subject, geographic, source, and type parameters, the user can require that the system delivers only relevant information.

Often such parameters can reduce a very large number of pieces of information to a much smaller number of pieces of information — a difference of orders of magnitude. Although there are now nearly half a million items in the INIS file, one can easily go to INIS and, by specifying search parameters quite precisely, select as few as 40 or, in some cases, as few as 4. The selective mechanism has the power to reduce a set by factors of the order of 10^5, and this is quite possible with our present technology. The subset, of course, has to be relevant to the user. You can measure the effectiveness of the system by the degree to which the selection works, whether the subset that is produced is relevant, whether there are items that are not relevant — "noise" that the system generates. If the information system contains relevant items and does not deliver these when the user puts in a command, the system's operation is ineffective. Thus there are methods for looking at an information system and seeing whether or not it is effective.

There are many different kinds of information systems handling different types of information. One can have information systems where the unit items are descriptions of projects: research projects, development projects. There are information systems where the units are descriptions of people and their talents, their capabilities to do particular types of work: for example, one can have an information system that is a register of consultants, an information system that describes institutions and the work that they do, the capacities they have for doing work for others. And finally, perhaps most importantly, there are information systems that identify documents, documents that themselves contain information. These are called bibliographic systems: INIS and AGRIS are bibliographic systems.

I will concentrate on the bibliographic systems, partly because they are more numerous and partly because they serve as a model of the whole family of such systems. An information system does not have to be computerized; there are information systems that operate by manual methods or by semiautomatic methods. There is no magic number of records above which a computer becomes necessary for operating an information system. I believe that if more than 10 000 items a year are added to the system, then computerization is almost essential. With fewer than 10 000 items a year, one may manage with semiautomatic or manual methods.

Now, I would like to mention a term that I do not particularly like but that belongs in the jargon of the business, and if I am not careful I will use it without defining it: "data base." The information that is put into a system is compiled and held by the system in some sequence depending on the design of the system. The description of a document in a bibliographic system can range from 600 to perhaps 5000 characters. Each set of characters that describes one document must be kept together in a "record." If, say, 100 000 records are properly sequenced in a file (for example, on magnetic tape or on disk), they constitute a "data base."

I would like now to give a bit of historical perspective to the types of data base that have been constructed in the past and that are being constructed today. In fact, I'd like to go all the way back to the late 19th century when this type of work began (when, of course, there were no computers). At that time the lead was taken essentially by academies of science or by scientific societies in particular disciplines. For example, one of the first information systems, *Chemical Abstracts*, was started by the American Chemical Society. In what is now the Soviet Union, the Academy of Sciences began the *Referativny Zhurnal*. In the early days it was the scientists themselves who took the lead; the methods were purely manual, and the information systems were built around a scientific discipline. There were services in chemistry, services in physics, and, later on, services in different aspects of engineering.

If one proceeds through the years into the interwar period or the period immediately following the Second World War, one finds that there was a shift and that the initiatives for new information systems were increasingly assumed by quasigovernmental bodies — especially by national research councils. In France, for example, the Conseil national de recherche scientifique developed the *Bullétin signalétique*. A little later — mostly during and after the Second World War — one begins to see the development of information systems whose scope is defined not by scientific disciplines but by economic activities. A good example is the information systems that were built by atomic energy commissions and were precursors of INIS. The U.S. National Aeronautics and Space Administration developed a space-science information system, and the U.S. National Agricultural Library, an agricultural information system. People began to realize that information had economic value and that it was worth investing public funds for the development of information systems. Now perhaps, the shift has gone even further: central agencies of government are taking responsibility for building information systems and for providing the funds and the investments that are needed. Sometimes, the initiative comes from the presidency of a republic, as in the case of Morocco, sometimes from a ministry of planning or of coordination as in the case of Mexico or Bolivia.

The shift in responsibility away from the scientific community and toward the central government agencies has been accompanied by a shift in the way the systems are defined. New systems are seldom now defined by discipline. The old systems — for biology, for chemistry, for physics — continue to meet important needs, but the new systems are not discipline-oriented. They are mission-oriented. The mission is some economic purpose — to promote the peaceful uses of atomic energy, to grow more food — which can be realized only if one collects information from many different disciplines; so the mission-oriented systems are multidisciplinary.

The most recent systems that have been started or proposed are ones that take economic and social development as the mission, and the purpose is

identifying and delivering information that can be useful in planning and in decision-making related to the development of a country. Hence they are concerned not only with scientific and technical information that may be important in decision-making but also with economic information, social information, legislative information.

So far, I have covered two words: "information systems." I would now like to bring in the word "cooperative." What is meant by "cooperative information systems"? Cooperative information systems are those that are produced and maintained by partner organizations. I think the move toward cooperative information systems comes from a realization that no one institution can do the whole job unless the subject is extremely narrowly defined. Language presents a problem: useful information appears in many different languages. It is true that a few languages predominate, but if one restricts oneself to those few languages, then one may miss a great deal of useful material. More important than the language problem, I believe, is the problem of how to find the information. In the old days, when the focus was on discipline-oriented information, one could find most of it in papers published in scientific journals. Scientists were interested in new discoveries, new increments to knowledge, and the scientific journals cover this information well — probably 90% or more of what is needed. But with the move toward mission-oriented systems where the concern is *application* of knowledge not only in science but in economic and social development, looking in the scientific journals is not enough. One must look in a lot of other places too. One needs to find the nonconventional documents — reports produced by different institutions, patents, theses, etc. Government documents, statements that are made by ministers, answers to questions in a parliament, all these items are important in the new context. The interest is not only in discoveries but also in experiences. The way in which a country applies knowledge to development projects may not be novel, but the result — be it positive, negative, or indifferent — is of value to people, particularly within the country itself. To find such information, one needs to search within the country. One cannot sit in London or Washington, D.C., or Moscow, and be able to get this kind of information from Mexico and Argentina and Kenya. Only the Mexicans, the Argentinians, and the Kenyans can find this information within their countries. Consequently one does need cooperation.

A cooperative information system is one where several, maybe many, institutions are involved in discovering information, reporting it so that it can be added to a data base, and then using the data base for their own purposes. Some moves have been made to make the older systems more cooperative. The American Chemical Society, for example, now has people working for it in many different countries to help find chemical information and report it for inclusion in the data base *Chemical Abstracts*. Making such a move brings into play technical factors that mean one must be a good deal more organized than when a system is operated by a single institution, particularly if the system is computerized. Computers are very unintelligent; they cannot interpret words that are misspelled or only to a very limited degree. One must apply rigorous standards in a computerized system; if it is a cooperative system then, of course, all the cooperating partners must apply the same standards. The standards must be prepared, they must be agreed, they must be written down. And then the participants must be trained; so a training program for the different participants is a first consideration. One must build in quality control to check what the participants are doing. And, obviously, if an institution is going to contribute input,

it is also going to want to have its say about the output. It wants to be sure that outputs are provided that are going to be useful to it as well as to the other organizations that are involved. So, along with cooperation for the preparation of input, one also has to build new management mechanisms that give a voice to each of the participants.

Now let me add the word "international" in front of the three terms already described. An "international cooperative information system": what is meant by that? In the context of this meeting, I think the adjective "international" must be applied in its full sense — cooperation among nations, cooperation among the governments of nations. There are strong incentives for governments to get together and organize this type of cooperation: but how do they do so? Governments do not want to build new international mechanisms unnecessarily; so, wherever possible, they will choose to use existing mechanisms. In general this means the international, intergovernmental agencies that make up the United Nations system. It is these agencies that are called upon to build the international cooperative information systems.

Why should governments be involved? Why not just let scientific societies handle the work? I think the reason that governments are becoming involved in this type of international cooperation is that governments are some of the most significant producers of information, particularly information that is clearly mission-oriented. Governments control the information they produce; they decide whether to make it available — to their citizens and to foreigners. Governments are more ready to do so if they can look forward to receiving information from other governments in exchange. And also governments are potentially the major users of information. If information systems are to be mission-oriented, used in the pursuit of goals in policymaking and decision-making, then the information must be delivered to government officials who will use it in their performance of these functions — and this can best be assured if governments are in fact the members of the community that cooperate to build the system. What are the major characteristics of a cooperative information system? There are significant common characteristics that have been embodied in INIS, that have been adopted by AGRIS, and that have been proposed for several other systems. The most important of these characteristics is what I call "the territorial formula": each member country is responsible for identifying the relevant information that is produced within its own territory — and *only* the information produced in its territory. Austria identifies the Austrian information; France identifies the French information; Canada identifies the Canadian information. And, as each country reports its information, this is added to the data base. The territorial formula provides something that is useful technically and that, I believe, is equitable from the point of view of the sharing of work among nations. It is useful technically because it eliminates any possibility of duplication. Canada will not report Austrian information; Austria will not report Canadian information. If one were to divide up responsibilities in some other way, by subject for example, and asked Canada to report information in one field and Austria to report information in another subject field, the separation would not be as precise because subjects overlap with each other and selection becomes a matter of intellectual judgment. So the territorial formula ensures that there is no duplication of information coming into the system. But it is also equitable: countries that have big programs have a lot of work to do, whereas countries that have small programs have a smaller amount of work to do, for the amount of information that a country produces is roughly proportional to the amount of

activity it has in the subject area. But then each country is entitled to receive the total data base. It is up to the management of the system to offer the data base in the forms in which it can best be used. In INIS and AGRIS the data base is offered both in printed form with indexes and on magnetic tape so that it can be searched on national computers. Each country can select from the total data base the information that is most important to it.

Moreover, the formula ensures that all the participants have a voice in the management of the system. There is no central authority arbitrarily saying, "it shall be so." What is necessary is the development of consultative mechanisms by which the participating countries can sit down with each other and agree on how a system is going to be developed. The rules are accepted by consensus. And this has big advantages for the participants: each knows that the rules will not be changed suddenly or arbitrarily. Rules will be changed only by consensus, and each participant has a voice in that consensus. So the system has a certain built-in stability, one that a more centralized operation might not have. Another benefit that has been developed is one that ensures that when payments must be made, these payments are in national currency. This contrasts with what applies in the case of centralized systems, for the products of which one must pay in the currency of the country where the system has been developed; this is not easy for many of the developing countries. On the other hand, if a system is managed by an international agency in the UN family, the first set of products come free of charge. If more are needed, then — as with all publications of the UN family whether on paper, on microfiche, or on magnetic tape — mechanisms exist that permit one to pay in local currency.

Finally, of course the central costs of such a system are borne on the budget of the intergovernmental agency that is acting as manager on behalf of all the countries. These are therefore shared among the participants according to the usual United Nations assessment formula. The rich countries pay more, and the poorer countries pay less. I submit that one cannot think of anything more equitable than that! In the declaration on a new international economic order, there are several statements that call for the sharing of knowledge among countries on a basis of equity. I believe that, in the international cooperative information system, the mechanism for sharing information among countries is on the basis of equity. I am hoping that, as the week proceeds, the participants here who are from developing countries will put their questions to those of us who have been concerned with the design and operation of such systems. I believe that we can demonstrate that it is a most equitable device for making knowledge available to all.

Now I would like to try to anticipate one of the questions that might be asked. Given a big data base with hundreds of thousands of documents described in it, some selection parameters so that out of that big data base one can select the dozens or the hundreds that may deal with a particular topic, what is available is still a lot of undigested material, really only an inventory. How can one digest this and select the documents, the facts, the figures, the opinions that guide in decision-making? And let me immediately admit that these big international cooperative information systems do not go that far; they do not digest. But I think one can also ask the question, should they digest? Because immediately one starts that process, there is another judgment being applied. Whose judgment will be acceptable? What biases are going to be put into the digestion process?

Should not the users do the digesting, in their own organizations — whether ministries or research institutions — so that they can be sure that no biases are put in

or, if biases are put in, at least they are the user's own biases? It is true, of course, that the top decision-maker cannot spend days reading hundreds of pages of information. But a top decision-maker has staff, and perhaps that is their chief responsibility — to digest information for the decision-maker.